AD_____

AWARD NUMBER: W81XWH-05-1-0283

TITLE: The Hygiene Hypothesis and Breast Cancer: A Novel Application of an Etiologic Theory for Allergies, Asthma, and Other Immune Disorders

PRINCIPAL INVESTIGATOR: Christina A. Clarke, Ph.D.

CONTRACTING ORGANIZATION: Northern California Cancer Center
Fremont, CA 94538

REPORT DATE: May 2009

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
1 May 2009	Final	1 May 2005 – 30 Apr 2009

4. TITLE AND SUBTITLE

he Hygiene Hypothesis and Breast Cancer: A Novel Application of an Etiologic Theory for Allergies, Asthma, and Other Immune Disorders

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-05-1-0283

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Christina A. Clarke, Ph.D.

E-Mail: tclarke@nccc.org

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Northern California Cancer Center
Fremont, CA 94538

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The "hygiene hypothesis", the idea that reduced exposure to important microbes, especially in childhood, impacts development of asthma and allergies, may have application to breast cancer. This research project aims to explore the hygiene hypothesis as it might relate to breast cancer development, thereby assessing its utility for more comprehensive future research. This research project aimed to interview a population-based series of Californian women recently diagnosed with breast cancer and a matched set of healthy control women as regards age-specific experiences relevant to microbial exposures. We interviewed by telephone 379 women aged 50-79 recently diagnosed with invasive breast cancer during the period 1/1/2004 and 9/31/2005 and 310 community women without breast cancer, matched on age and race and identified through mailing lists. After consideration of established risk factors for breast cancer, we found significantly protective associations worthy of further analysis for several categories of exposures including 1) school exposures, including attendance at preschool and kindergarten, and ever living at a boarding school where one lived in a dorm. 2) living within ½ mile of barns or stables; and 3) current consumption of lactobacilli-containing supplements. These findings are now being followed up in a separate application to the National Cancer Institute.

15. SUBJECT TERMS
immunologic exposures, infectious exposures, early life exposures, socioeconomic status, population-based, multiethnic, case-control study

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	33	19b. TELEPHONE NUMBER *(include area code)*

Table of Contents

Introduction

Breast cancer is the most commonly diagnosed cancer in women worldwide, and incidence and mortality rates have increased substantially over the past 50 years. Reasons for these increases are not entirely clear, because breast cancer causes remain incompletely understood. In the absence of means of primary prevention for breast cancer, partial understanding of its causation compels research into new etiologic hypotheses. Identification of novel hypotheses with promise for detailed etiologic investigation should take into consideration the established features of breast cancer epidemiology. A group of factors meeting these criteria are those mediating reduced exposure to microbes, especially in early life. Microbial exposures in early life are thought to be critical to the development of a robust immune system and have been well studied in the etiologies of allergies, asthma, autoimmune disease and other disorders of immune function. This research has led to the “hygiene hypothesis”, the idea that reduced or delayed exposures to important microbial inputs hamper the development of a healthy immune system in early life and the maintenance of such a system in adult life, which in turn increases vulnerability to the development of chronic diseases. This research project represented a preliminary effort to flesh out the hygiene hypothesis as it might relate to breast cancer development, thereby assessing its utility for more comprehensive future research. We interviewed by telephone 379 women aged 50-79 recently diagnosed with invasive breast cancer and 378 community women without breast cancer, matched on age and race and identified through newly developed address-based sampling procedures (see below). Eligible for inclusion as cases were all women aged 50-79 diagnosed with histologically confirmed, primary invasive BC during the period 1/1/2004-9/31/2005 while resident in San Mateo, Santa Cruz, San Benito, or Monterey counties, California, as reported to the Greater Bay Area Cancer Registry. Study controls were women without breast cancer living in the same counties, selected to be frequency-matched to cases on race/ethnicity and 5-year age category. Control subjects were ascertained using a novel, address-based sampling procedure.

Body

This project was a preliminary study to investigate possible associations of proxy measures of microbial exposures across the lifespan as they might associate with breast cancer risk, with the ultimate goal of identifying exposures for further, more intensive research. At the end of our study period, which included a no-cost extension year, we had accomplished all of the tasks as laid out in our approved Statement of Work. We experienced significant delays at several points in the study, specifically those involving obtaining local and DOD approvals as well as tasks involving subject interviewing. Below, we provide detail on each task in the Statement of Work and its final disposition. We also provide a full summary of the scientific accomplishments of this study, the reportable events, and consider the consistency of these findings with that of previously reported findings.

Statement of Work task summaries

Task 1. *Develop structured questionnaire, including appropriate questions about microbial exposures for breast cancer by translating concepts from existing hygiene hypothesis literature and incorporating questions from questionnaires from breast cancer studies and a Hodgkin's disease study (months 1-3)*

a. *Compile comprehensive list of topics from hygiene hypothesis literature*

   We carefully reviewed the hygiene hypothesis literature and compiled a comprehensive list of topics to be included in the questionnaire (Year 1).

b. *Obtain language for some questions from Dr. Liu and other authors*

   We obtained language for some questions from Dr. Liu and other authors, e.g. allergy section of the questionnaire (Year 1).

c. *Pilot test for appropriateness for women aged 50 and older*

   We tested the questionnaire for research question appropriateness for women aged 50 and older. This entailed consideration of cohort-specific experiences in formulating and editing questions (Year 1).

d. *Pilot test for variation in concept by ethnicity*

   We pilot tested the questionnaires among women of white, Asian, and Hispanic ethnicities and incorporated input from them regarding answer choices and terminology. For example, we added different housing types to a question about childhood housing when an Asian woman responded that she grew up in a barracks (Year 1).

e. *Pilot test reliability when asked of same person*

   We tested reliability when asked of same person by testing slightly revised versions of the questionnaire on the same women, then comparing the answers to see if the answer was comparable (Year 1).

Task 2. *Finalize study documents, obtain needed approvals and complete other preparations for study commencement (months 4-9)*

a. *Finalize questionnaire and verbal consent scripts*

The final questionnaires (Appendix A) and consent form in English (Appendix B) and Spanish (Appendix C) are attached. Briefly, the questionnaire inquired about

- Detailed birthplace and immigration status of self and parents
- Detailed educational status of self and parents
- Natural or c-section delivery
- Detailed exposure to siblings (full, half, adopted) and other children before age 12
- Attendance at daycare/preschool before age 5; at kindergarten; at boarding school
- Home environment (socioeconomic markers, farm vs. urban, furry pets in home, within 0.5 miles of stables, at ages 6 months, 3, 6, 12, 30 and now
- Exposure to children in adulthood (own, adopted, raised, teaching, daycare)
- Age at infection with parasites, Lyme disease, chickenpox, mastitis, chlamydia, recurrent urinary tract or gum infections, other infections
- Detailed history of asthma and allergy, including allergen types
- History of tonsillectomy, appendectomy, splenectomy
- Current consumption of fermented or probiotic foods (e.g., yogurt, kim chee)
- Lifetime number of insect stings

b. Translate questionnaire to Spanish

The questionnaire and other relevant study documents were translated into Spanish (Appendix D)

c. Obtain Institutional Review Board approvals

After many months of communications, we received final approvals for this research project from the NCCC Institutional Review Board on 9/11/07 and from the HSRRB on 7/26/06.

d. Create study tracking system

In Year 1, we created the study tracking system using Microsoft Access software. The system includes capacity for Computer Assisted Telephone Interviewing (CATI) to improve telephone interviewers' efficiency in data collection. In Year 2, this system was updated to include capacity for Spanish language interviewing and control subject frequency-matching.

e. Hire and train interviewers

In Year 2, we hired and trained 3 interviewers, one of whom was bilingual in Spanish.

Task 3. Recruit a selection of women recently diagnosed with invasive breast cancer, and age- and race-matched healthy women and interview them about hygiene-hypothesis-

relevant exposures as well as established breast cancer risk factors, using study questionnaire (months 9-27)

We originally planned to begin interviewing subjects by month 9 of Year 1 (March 2006). However, it took longer than we anticipated receiving NCCC IRB and HSRRB clearances. Upon completion of the pilot testing and planning process, we made several changes to the statement of work from its original form, include reducing sample size from 1050 women to 1000, modifying the means of control selection from random-digit dialing to an address-based sampling procedure, dropping the life calendar from the subject mailing, and including a saliva specimen retrieval kit with the mailing, as described below. Dr. Carole Christian, our Army Contracting Officer Representative, confirmed via email on 5/30/06 that these changes were not significant enough to warrant a formal change of SOW.

a. *Obtain listing of eligible cases from population-based Greater Bay Area Cancer Registry*

We obtained all relevant approvals and clearances from the Greater Bay Area Cancer Registry and have received listings of all eligible cases. In light of the unexpected delays in obtaining human subjects approvals as described above, we altered the dates of diagnosis from our original proposal in order to have the greatest chances of contacting and interviewing patients recently diagnosed with breast cancer. Instead of trying to recruit women diagnosed 1/1/2003-7/30/2004, we instead requested listings for women diagnosed between 10/1/2004 and 9/31/2005. By the end of the study, we had received listings for 743 breast cancer patients meeting our age and residency requirements.

b. *Establish random-digit dialing (RDD) procedures to ascertain control subjects and conduct RDD*

After consulting with study collaborators, particularly co-investigator Dr. Pamela Horn-Ross who is experienced in the design and conduct of breast cancer case-control studies and RDD, we decided to modify the methodology used to ascertain control subjects using a novel, address-based sampling procedure. This procedure follows many of the principles of traditional RDD but has the additional advantage of allowing for mail, telephone, and personal modes of recruitment. In addition, it provides a known sampling frame, which is no longer possible with RDD. This methodology is described in detail in the study protocol and summarized briefly below. In February 2007, we purchased a “saturation list” mailing address list from Marketing Systems Group. The list represented a n=10,000 random sample of all US mail-deliverable addresses for San Mateo, Santa Cruz, San Benito, and Monterey counties. Using mailing lists based on residency offers a way to sample the same general population from which the breast cancer cases occur, a fundamental principle of control selection. Introductory letters containing $1 bills were mailed to each address selected. These letters request that recipients call a toll-free line or use email to enumerate their household. Women meeting our selection criteria (female sex, aged 50-79, no prior history of breast cancer) were frequency-matched to cases on five-year age group and race/ethnicity. For households that do not respond to one of the modes within a two-week time frame, we utilized Internet search databases to try and identify a working phone number for the candidate

control address. For candidate control addresses for which we can find a phone number (estimated to be 50%), a trained interviewer telephone to attempt to recruit eligible women. We initiated control ascertainment in April 2007 and continued through our one-year no-cost extension. Ultimately, we identified and interviewed 378 eligible control subjects.

c. Mail letters to physicians to ascertain contraindications to contact

We sent letters to physicians prior to contacting all patients (Years 1-3).

d. Mail letters of invitation to subjects

At the end of the study, we had sent letters of invitation to all 743 breast cancer patients and to 5500 potential control households.

e. Telephone subjects to confirm participation

At the end of the study, our interviewers had called 743 breast cancer patients and 3197 potential control households to invite eligible women to participate. We attempted to locate telephone numbers for all sampled control households using internet-based proprietary "skip tracer" resources but ultimately could not identify a working number for many of these.

f. Mail life calendar and informed consent guide to subjects

We learned from pilot testing the pre-interview that a life calendar substantially slowed the pace of the interview and did not seem to substantially aid subject recall, thus we decided to drop it from the pre-interview packet to be mailed to subjects. We also decided to include with the informed consent documentation in the pre-interview packet a saliva specimen to be mailed back by each participant.

g. Interview subjects by telephone

At the end of the study, our interviewers had completed interviews with 757 subjects (n=379 cases and n=378 controls) by telephone. Our response rates were considerably lower than we had anticipated, as detailed below.

<u>Case participation rates:</u> Of n=743 cases identified, we were able to successfully interview 51% (n=379). Some of these cases were not deemed eligible for the study upon further contact. 2 cases had their physician disallow contact from our study. We also learned from trying to contact cases that 3% (n=20) were recently deceased. A further 6% of cases (n=44) were ultimately deemed ineligible for the following reasons: being too ill to participate in the interview (n=12), not speaking fluent English or Spanish (n=24), or having comprehension problems/senility (n=8). Of the remaining potentially eligible cases (n=675), the numbers refusing to participate in the study were higher than we would have anticipated, including 3.5% who used the opt-out box on the initial response form (n=26) and 15% "hard-refusing" (n=109) on telephone contact. An additional 2% of cases (n=15) refused to participate because of their concerns about other DOD-funded non-research activities. By the

end of the data collection period, 101 breast cancer cases "soft-refused" or "passively refused", meaning they did not respond to at least 10 efforts to contact them by mail or telephone, bringing our estimate of the total refusal rate (hard+soft) to 33%, and our estimated response rate among cases to 66%.

Control participation rates: Calculation of control participation rates is considerably more difficult than that for cases, because failure of a sample household to respond to the initial mailed letter of invitation generally did not reveal any information about the eligibility of any of the household residents (e.g a woman aged 50-79 with no history of breast cancer). Of the 5500 control households approached by mail to participate in the study, 2481 (45%) households ultimately "soft-refused" by not responding to at least 10 efforts to contact them by mail or telephone, if a telephone number could be located for the sampled address. A further 996 households could be reached by telephone, but refused to offer any information about the eligibility of any of the residents. Some households that did offer information about eligibility were deemed ineligible on the basis of not having an eligible household member of a woman aged 50-79 at the time (n=1042) or "hard-refused" to participate in the actual interview (n=292) including some women who refused to participate refused to participate because of their concerns about other DOD-funded non-research activities (n=14. Some households with women eligible on the basis of age were later deemed ineligible because the potential control had had breast cancer (n=51), suggested that they were too ill to participate in the interview (n=4), did not speak fluent enough English or Spanish (n=45), had comprehension problems/senility (n=19). Of the control households that we were able to establish as having an eligible member, we recruited and successfully interviewed 378 women.

h. Send subjects thank you note and compensation

We sent all participants thank you letters along with $10 compensation (check or gift card) for participating in the study.

i. Call back subjects to resolve discrepancies

In a very limited number (n=4) of circumstances, interviewers needed to call back subjects after completion of interview to resolve discrepancies on particular questionnaire items or codes. The use of a computerized assisted telephone interview system greatly reduced the need for these kinds of calls because 1) the computer would not allow the interviewer to enter codes outside the range of acceptable codes and 2) it provided the interviewer with several opportunities to record information needed to resolve discrepancies.

j. Enter and clean data to create analytic database

The database was cleaned and each variable subjected to several consistency checks upon completion of data collection.

Task 4. Preliminary analyses: Evaluate whether risk of breast cancer diagnosed in women aged 50 years or older is associated with "hygiene hypothesis"-relevant exposures independently of reproductive characteristics and other established BC risk factors (Specific Aim 1) and assess whether associations could be

limited to select demographic or tumor groups (Months 28-32)

a. *Compare distributions of these characteristics between cases and controls*

Preliminary univariate comparisons of cases and controls suggested several differences in some but not all established risk factors for breast cancer (e.g., hormone therapy use history, biopsy for benign breast cancer) and novel microbial risk factors. A comprehensive list of the associations, including statistical significance testing with all of the microbial risk factors assessed in the questionnaire are shown in **Table 1** below. For most microbial exposures assessed, we did not observe significant differences in distributions between cases and controls, including birthplace or immigration status of self or parents; educational status of self or parents; mode of delivery; exposure to siblings (full, half, adopted) or other children before age 12; exposure to children (own, grandchildren, teaching) in adulthood; infectious disease history; mastitis history; asthma and allergy history; history of tonsillectomy, appendectomy, or splenectomy; current consumption of yogurt or kimchee; or number of lifetime insect stings. However, suggestively protective associations worthy of further analysis were observed for several categories of exposures 1) school exposures, including attendance at preschool and kindergarten, and ever living at a boarding school where one lived in a dorm. (Having ever been home schooled was not associated with breast cancer, however) 2) living within ½ mile of barns or stables, but only at particular ages, including ages 12, 30 and the reference year; and 3) consumption of lactobacillus or other probiotic-containing supplements.

b. *Estimate relative risk by calculating odds ratios for suggestively associated risk factors*

We used multiple logistic regression to calculate odds ratios (OR) for the suggestively associated risk factors described above. These odds ratios and associated 95% confidence intervals (CI)are presented in **Table 2** below.

c. *Adjust these associations for possible confounders*

We used multiple logistic regression to adjust the suggestively associated factors for their association with breast cancer case status in two ways 1) for age and race/ethnicity only and 2) for all other established risk factors for breast cancer observed in this particular population (e.g., possible confounders). The factors constituting the set of confounders included first degree family history of breast cancer, current consumption of 10 or more alcoholic beverages per week, and current hormone replacement therapy use (none, current estrogen only, current combined estrogen-progestin use). These adjusted OR are shown in **Table 2** below. Most of the suggestive associations remained statistically significant after adjustment for all confounders. Never having attended kindergarten was negatively associated with breast cancer (adjusted OR: 1.52; 95% CI: 1.06-2.18). Never having attended preschool had an OR of similar magnitude, but this association was of marginal significance (adjusted OR: 1.47; 95% CI:

0.93-2.34). Never having lived in a boarding school where one lived in a dormitory with other students was associated with almost a doubling of risk (adjusted OR: 1.91; 95% CI: 1.11-3.29). Not living within ½ mile of barns or stables showed a distinct age-specific pattern, whereby exposure at ages 6 months, and 5 years was not statistically significantly related, but exposure at age 12 was associated with elevated risk (adjusted OR: 1.54; 95% CI 1.08-2.19) as was exposure at age 30 (adjusted OR: 1.57; 95% CI 1.05-2.33). Not living near barns and stables at the reference year was of marginal significance but of comparable effect size (adjusted OR: 1.42; 95% CI: 0.98-2.04). Consumption of probiotic supplements in the last year was not associated with breast cancer risk after adjustment for age and race/ethnicity.

d. Explore possible effect modification by race/ethnicity and tumor characteristics

We stratified our analyses of the suggestive associations by race/ethnicity and tumor characteristics in order to explore any possible effect modification by these factors. Numbers of cases by race/ethnicity were limited by specific ethnic group, with 312 cases (82%) and 309 controls (82%) being white, non-Hispanic, 26 cases (7%) and 34 controls (9%) Hispanic, and other groups constituting fewer than 10 each of cases and controls. Thus, we carried out multiple logistic regression for two groups, "white, non-Hispanic" and "non-white", as shown in **Table 3** below. Association effect directions and sizes were generally similar between the two groups, with a few exceptions. The association with boarding school attendance was shown to be limited to white, non-Hispanics only, and remained statistically significant with an adjusted OR: 2.13; (95% CI: 1.16-3.90). The association with use of lactobacillus or other probiotic containing supplements was not significant in either race group, but showed opposite directions of effect, with a suggestively protective effect in whites but a deleterious effect in non-whites. We also explored heterogeneity of suggestive associations according to tumor characteristics including estrogen receptor (ER) status (ER positive and unknown vs. ER negative) and tumor histology (ductal vs. non-ductal). **Table 4** shows odds ratios for ER positive breast cancers as compared to all controls. Results were nearly identical to those observed for all breast cancers combined, although several associations that were significant in the full dataset were of marginal significance in this smaller subset.

e. Assess selection bias and consider influence on results

We successfully geocoded and linked all street addresses for cases (obtained from cancer registry listings) and controls (obtained from marketing company mailing list) to the 2000 US census to obtain neighborhood-specific information regarding socioeconomic characteristics. This allowed us to examine odds ratios for breast cancer according to neighborhood characteristics for two groups: 1) participants in the study (379 cases, 378 controls) and the larger pool of presumably eligible subjects (780 cases, 3550 control households) with the theory being that in the absence of selection bias, associations with neighborhood characteristics should be comparable. **Table 5** shows odds ratios for neighborhood characteristics for the two groups and a calculated measure of bias between the two comparisons. The participating groups are shown to

derive from higher socioeconomic neighborhoods with respect to percent of residents with higher incomes and higher education. This bias would be expected to result in underestimates of the associations we observed, as the controls are too similar to the cases with respect to markers of higher socioeconomic status.

Key Research Accomplishments

- We successfully identified several key categories of proxy measures of microbial exposures with protective associations with breast cancer risk, thereby meeting the major goal of this exploratory study, which was to screen the exposures reported as promising from the asthma/allergy literature to identify exposures for further, more intensive breast cancer etiologic research.

- Specifically, we identified as promising exposures 1) ages at first school attendance and characteristics of school relevant to intensity of other exposures to children (e.g., boarding school where one lived in a dormitory); 2) age-specific residential location nearby barns or stables where livestock are kept; 3) current consumption of lactobacillus supplements.

- Used these findings as preliminary data to apply to the National Cancer Institute for a new study to further investigate these exposures and genetic variation in innate immunity genes that might modify their mechanism of action.

- Developed and successfully implemented a novel, address-based methodology for identifying and recruiting population-based control subjects using mixed mail/telephone modes of recruitment. This methodology allows for precise quantification of Census-based neighborhood differences between recruited participants and households who did not respond to the invitation to participate. This methodology is being used in applications for future case-control studies.

Reportable Outcomes

The following represent reportable scientific outcomes of this project:

Scientific presentation: Clarke CA, Horn-Ross PL, Glaser SL. Microbial Exposures And Risk Of Postmenopausal Breast Cancer: A Population-Based Case-Control Study (poster presentation). Department of Defense Era of Hope biannual meeting, Baltimore, MD, June 2008.

Grant application using this data as preliminary data: Microbial burden, innate immunity, and risk of hormone-sensitive breast cancer. R01 application submitted by Dr. Christina Clarke as Principal Investigator to the National Cancer Institute, submitted June 5, 2009.

Data resource: the database created from this project represents a rich resource for further examining associations of environmental exposures and breast cancer, suitable for pooling in consortial analyses. The data resource includes interview information, saliva-obtained DNA specimens, cancer registry-derived demographic and tumor information, and has the capacity for up-to-date vital status and survival time information from continued linkage with the cancer registry.

Relationship of these findings with that of previously reported findings:

To our knowledge, there have not yet been any findings published that address early life microbial exposures and breast cancer risk. Our findings of protective effects of living near barns and stables are consistent with some prior findings from a North Carolina case-control study suggesting that breast cancer risk was lower among women who had lived or worked on farms, with a dose-response effect for duration of farming with odds ratios of 0.7 (95% CI 0.5-1.1), and 0.6 (95% CI 0.4-0.9) for 18-23 and over 23 years of farming, respectively, relative to women who had never farmed[1]

Our findings also are comparable to some other preliminary and unpublished data collected by Dr. Clarke as part of her work with the California Teachers Study, a prospective cohort of over 133,000 California Teachers and school administrators. Participants responding to a fourth follow-up questionnaire administered in 2005/2006 were asked about five exposures relevant to the hygiene hypothesis, including characteristics of their home environment at ages 6 months, 3 years, 5 years, 12 years, 30 years, and "now" (rural, small town, suburb or urban area; within half a mile of barns or stables where horses, cows, pigs, or other hoofed animals were kept; number of siblings or other people sleeping in the same bedroom, cat or dog living inside the home) and regular attendance (at least 30 times/year) of preschool; kindergarten or other regular gathering of at least 4 other children (ages 6 months, 3 and 5 only). We detected similar protective associations to those observed here for preschool attendance, kindergarten attendance, and living within a half-mile of barns or stables, but at ages 6 months and 3 years as opposed to older ages (30 and now).

Conclusions

This project has provided important leads as to the role of early life, immunocalibrating exposures in protecting against the development breast cancer. 1) ages at first school attendance and characteristics of school relevant to intensity of other exposures to children (e.g., boarding school where one lived in a dormitory); 2) age-specific residential location nearby barns or stables where livestock are kept; 3) consumption of lactobacillus supplements These leads now are being followed in two ways: 1) a larger R01 application to the National Cancer Institute submitted June 2009 and 2) a new collaboration of Dr. Clarke with a fellow DOD grantee, Dr. Susan Erdman of the Massachusetts Institute of Technology, who also applied in June 2009 for NIH funding to look at gut bacteria and mammary tumors in mouse models. Dr. Clarke is serving as a consultant on Dr. Erdman's application and Dr. Erdman is a consultant on Dr. Clarke's application. These projects and future work in human subjects will be able to better understand these associations by requesting more detail regarding the suggestively associated exposures, and by examining genetic variation in toll-like receptors and other functional aspects of the innate immune system that mediate microbial exposures.

Our research may ultimately have impact on breast cancer prevention. Our observations of a possibly protective influence of living near barns and stables may help to explain urban/rural, socioeconomic, or racial/ethnic variation in breast cancer incidence, which may help to alleviate the marked disparities in breast cancer occurrence observed by these parameters. With respect to urban/rural differences in breast cancer, our observations of protective effects of living near barns and stables firmly underscore the importance of continuing to look for aspects of farming and rural lifestyles that are protective against breast cancer development, as opposed to the predominant research direction of trying to determine what kinds of exposures common in urban environments are associated with increased risk of breast cancer. If specific types and timing of microbial exposures can be subsequently corroborated and determined to be causal for breast cancer using future research, these observations could ultimately lead to feasible primary prevention efforts for young women, perhaps through vaccination with harmless surrogates of important microbial exposures.

List of personnel receiving pay from this project

Christina A. Clarke, Ph.D, principal investigator
Pamela Horn-Ross, Ph.D, coinvestigator

Sally L. Glaser, Ph.D, coinvestigator
Sarah Shema, MS, biostatistician
Lily Huynh, project coordinator
Trisha Harasty Weeks, program manager
Rekha Subramanyan, programmer/analyst
Zinnia Loya, interviewer
Kalliope Bellas, interviewer
Jolyn Smith, interviewer
Sarah Aroner, research assistant
Dawn Beahm, research assistant
Tram Nguyen, project assistant

References

1. Duell EJ, Millikan RC, Savitz DA, et al. A population-based case-control study of farming and breast cancer in North Carolina. Epidemiology 2000;11:523-31.

Appendices and Supporting Data

Table 1: Distributions between female breast cancer cases (n=379) and controls (n=310) for all characteristics, including established risk factors for breast cancer and novel proxy measures of microbial exposures as gleaned from the "hygiene hypothesis" literature for asthma and allergies, Northern California, 2006-08.

Table 2: Odds ratios (OR) for breast cancer for suggestively associated proxy measures of microbial exposures as ascertained from cases (n=379) and controls (n=310), Northern California, 2006-08.

Table 3: Multivariately adjusted odds ratios (OR) for breast cancer and 95% confidence intervals (CI) for suggestively associated proxy measures of microbial exposures considered separately by race/ethnicity of participants, Northern California, 2006-08.

Table 4: Multivariately adjusted odds ratios (OR) for breast cancer and 95% confidence intervals (CI) for suggestively associated proxy measures of microbial exposures considered only for estrogen-receptor positive breast cancer (n=364), Northern California, 2006-08.

Table 5: Odds ratios for neighborhood characteristics and breast cancer among respondents in the study (379 cases, 378 controls) and the larger pool of presumably eligible subjects (780 cases, 3550 control households), including calculation of % difference between groups.

TABLES OF STUDY RESULTS

Table 1: Distributions between female breast cancer cases (n=379) and controls (n=378) for all characteristics, including established risk factors for breast cancer and novel proxy measures of microbial exposures as gleaned from the "hygiene hypothesis" literature for asthma and allergies, Northern California, 2006-08.

	Controls (n=378)			Cases (n=379)	
	N	%		N	%
Lifetime physical activity on job					
Mostly sedentary	52	13.76		71	18.73
Lightly active	69	18.25		67	17.68
Moderately active	187	49.47		165	43.54
Highly active	70	18.52		76	20.05
p value			0.2049		
Lifetime physical activity not on job					
Mostly sedentary	24	6.35		32	8.44
Lightly active	85	22.49		76	20.05
Moderately active	215	56.88		217	57.26
Highly active	54	14.29		54	14.25
p value			0.6472		
Ever had breast biopsy that showed benign breast disease					
Yes	**82**	**21.69**		**109**	**28.99**
No	**296**	**78.31**		**267**	**71.01**
p value			**0.0213**		
Father ever diagnosed with cancer					
Yes	135	35.71		117	30.87
No	234	61.90		244	64.38
Unknown	9	2.38		18	4.75
p value			0.1057		
Mother ever diagnosed with cancer					
Yes	124	32.80		126	33.25
No	248	65.61		248	65.44
Unknown	6	1.59		5	1.32
p value			0.9486		
Marital status					
Single	33	8.73		20	5.29
Married	236	62.43		235	62.17
Widowed	36	9.52		51	13.49
Separated/ Divorced	73	19.31		72	19.05
p value			0.1226		
Highest educational level obtained					
None	1	0.26		0	0.00
Grade school	13	3.44		11	2.90
High school	71	18.78		76	20.05
College	185	48.94		208	54.88
Graduate wk	108	28.57		84	22.16
p value			0.2242		
Self-reported race/ ethnicity					
White	309	81.75		312	82.32
Black	8	2.12		4	1.06

Hispanic	34	8.99		26	6.86
Chinese	7	1.85		10	2.64
Japanese	2	0.53		1	0.26
Filipina	5	1.32		11	2.90
Korean	1	0.26		0	0.00
South Asian	1	0.26		2	0.53
Pacific Islander	1	0.26		2	0.53
Native AM/ Alaskan	1	0.26		2	0.53
Mixed Race	9	2.38		7	1.85
Unknown	0	0.00		2	0.53
p value			0.5506		
Foreign Born					
US Born	312	82.54		308	81.27
Foreign Born	66	17.46		71	18.73
p value			0.6492		
Mode of one's own birth					
Vaginally	363	96.80		364	97.07
Cesarean	12	3.20		11	2.93
p value			0.8323		
Before age 5, went to day care					
Yes	20	5.31		23	6.07
No	357	94.69		356	93.93
p value			0.6504		
Attended preschool					
Yes	**54**	**14.29**		**37**	**9.76**
No	**324**	**85.71**		**342**	**90.24**
p value			**0.0557**		
Attended kindergarten					
Yes	**305**	**80.69**		**273**	**72.03**
No	**73**	**19.31**		**106**	**27.97**
p value			**0.0051**		
Before 18, schooled at home					
Yes	7	1.85		10	2.64
No	371	98.15		369	97.36
p value			0.4654		
Before 18, boarding school where lived in dorm					
Yes	**40**	**10.58**		**24**	**6.35**
No	**338**	**89.42**		**354**	**93.65**
p value			**0.0366**		
Father born in US					
Yes	267	70.63		261	68.87
No	110	29.10		116	30.61
Unknown	1	0.26		2	0.53
p value			0.756		
Mother born in US					
Yes	274	72.49		285	75.20
No	104	27.51		93	24.54
Unknown	0	0.00		1	0.26

p value			0.4006		
Father's highest level of education					
None	**6**	**1.59**		**5**	**1.32**
Grade school	**61**	**16.14**		**79**	**20.84**
High school	**133**	**35.19**		**135**	**35.62**
College	**117**	**30.95**		**87**	**22.96**
Graduate wk	**49**	**12.96**		**40**	**10.55**
Don't know	**12**	**3.17**		**33**	**8.71**
p value			**0.0036**		
Mother's highest level of education					
None	7	1.85		7	1.85
Grade school	55	14.55		61	16.14
High school	177	46.83		199	52.65
College	110	29.10		89	23.54
Graduate wk	20	5.29		11	2.91
Don't know	9	2.38		11	2.91
p value			0.2499		
Have any siblings					
Yes	357	94.44		351	92.86
No	21	5.56		27	7.14
p value			0.3708		
Average number of days/month in last year reported eating yogurt					
None	108	28.57		109	28.76
1-5 day/ mo	120	31.75		142	37.47
6-10 days/ mo	55	14.55		49	12.93
11+ days/ mo	95	25.13		79	20.84
p value			0.2996		
Average number of days/month in last year reported eating kimchee or other pickled vegetables					
None	343	90.74		331	87.34
1+ days	35	9.26		48	12.66
p value			0.1337		
Average number of days/month in last year reported taking lactobacillus or probiotic supplements					
None	**350**	**92.59**		**362**	**95.77**
1 + days	**28**	**7.41**		**16**	**4.23**
p value			**0.0623**		
Number of lifetime bee or wasp stings					
None	49	12.96		61	16.09
1 time	63	16.67		76	20.05
2 times	71	18.78		76	20.05
3 times	47	12.43		48	12.66
4-5 times	58	15.34		40	10.55
6+ times	90	23.81		78	20.58
p value			0.2307		
In the last five years, how covered by health insurance					
Covered all the time	351	92.86		347	91.56

Covered some of the time	17	4.50		14	3.69
No	10	2.65		18	4.75
p value			0.2729		
In the last five years, how many mammograms received					
0	**21**	**5.56**		**27**	**7.12**
1	**20**	**5.29**		**24**	**6.33**
2	**38**	**10.05**		**30**	**7.92**
3	**55**	**14.55**		**34**	**8.97**
4	**36**	**9.52**		**21**	**5.54**
5	**201**	**53.17**		**226**	**59.63**
6	**1**	**0.26**		**6**	**1.58**
7	**4**	**1.06**		**5**	**1.32**
8	**0**	**0.00**		**3**	**0.79**
10	**2**	**0.53**		**2**	**0.53**
17	**0**	**0.00**		**1**	**0.26**
p value			**0.0283**		
In the last five years, how many pelvic exams/pap smears received					
0	44	11.64		55	14.55
1	31	8.20		20	5.29
2	47	12.43		33	8.73
3	30	7.94		33	8.73
4	34	8.99		24	6.35
5	188	49.74		210	55.56
7	1	0.26		0	0.00
8	2	0.53		1	0.26
10	1	0.26		2	0.53
p value			0.2136		
At age 6 months, lived in rented house/ apartment					
Yes	187	50.00		198	52.94
No	187	50.00		176	47.06
p value			0.421		
At age 6 months, type of area lived					
Farm	24	6.37		29	7.71
Rural	35	9.28		26	6.91
Town	109	28.91		108	28.72
Suburb	62	16.45		64	17.02
Urban	147	38.99		149	39.63
p value			0.7637		
At age 6 months, housing had indoor plumbing					
Yes	342	92.18		334	90.51
No	29	7.82		35	9.49
p value			0.4195		
At age 6 months, self-ranked household financial status (1-5: highest)					
1 (lowest)	**13**	**3.48**		**28**	**7.65**
2	**101**	**27.01**		**95**	**25.96**
3	**213**	**56.95**		**178**	**48.63**
4	**33**	**8.82**		**51**	**13.93**
5 (highest)	**14**	**3.74**		**14**	**3.83**

p value			**0.0135**		
At age 6 months, how many others shared bedroom					
0	128	35.96		119	34.59
1	104	29.21		90	26.16
2	95	26.69		106	30.81
3+	29	8.15		29	8.43
p value			0.6292		
At age 6 months, how many cats/dogs/furry pets lived in the home					
Yes	13	3.48		10	2.70
No	361	96.52		360	97.30
p value			0.5424		
At age 6 months, house was within 1/2 mile of barns and stables where livestock were kept					
Yes	81	22.19		80	21.86
No	284	77.81		286	78.14
p value			0.9133		
At age 5, lived in rented house/ apartment					
Yes	147	38.89		148	39.26
No	231	61.11		229	60.74
p value			0.9174		
At age 5, type of area lived					
Farm	22	5.82		30	7.94
Rural	32	8.47		32	8.47
Town	109	28.84		105	27.78
Suburb	96	25.40		88	23.28
Urban	119	31.48		123	32.54
p value			0.7872		
At age 5, housing had indoor plumbing					
Yes	354	93.65		350	92.84
No	24	6.35		27	7.16
p value			0.6564		
At age 5, self-ranked household financial status					
1	**6**	**1.60**		**19**	**5.08**
2	**66**	**17.55**		**74**	**19.79**
3	**217**	**57.71**		**194**	**51.87**
4	**71**	**18.88**		**72**	**19.25**
5 (highest)	**16**	**4.26**		**15**	**4.01**
p value			**0.0737**		
At age 5, how many others shared bedroom					
0	118	31.38		121	32.18
1	172	45.74		162	43.09
2	51	13.56		58	15.43
3+	35	9.31		35	9.31
p value			0.8527		
At age 5 how many cats/dogs/furry pets lived in the home					

Yes	**20**	**5.29**		**10**	**2.65**
No	**358**	**94.71**		**368**	**97.35**
p value			**0.0625**		
At age 5, house was within 1/2 mile of barns and stables where livestock were kept					
Yes	85	22.67		89	23.61
No	290	77.33		288	76.39
p value			0.7597		
At age 12, lived in rented house/ apartment					
Yes	103	27.32		101	26.72
No	274	72.68		277	73.28
p value			0.8524		
At age 12, type of area lived in					
Farm	18	4.76		24	6.35
Rural	37	9.79		31	8.20
Town	104	27.51		105	27.78
Suburb	117	30.95		114	30.16
Urban	102	26.98		104	27.51
p value			0.8355		
At age 12, housing had indoor plumbing					
Yes	362	96.02		359	94.97
No	15	3.98		19	5.03
p value			0.4876		
At age 12, self-ranked household financial status					
1	8	2.12		6	1.60
2	46	12.20		44	11.70
3	200	53.05		206	54.79
4	99	26.26		98	26.06
5 (highest)	24	6.37		22	5.85
p value			0.9726		
At age 12, how many others shared bedroom					
0	185	48.94		171	45.24
1	128	33.86		139	36.77
2	29	7.67		40	10.58
3+	36	9.52		28	7.41
p value			0.2889		
At age 12, did cats/dogs/furry pets live in the home					
Yes	**46**	**12.17**		**29**	**7.67**
No	**332**	**87.83**		**349**	**92.33**
p value			**0.0386**		
At age 12, house was within 1/2 mile of barns and stables where livestock were kept					
Yes	**101**	**26.72**		**73**	**19.36**
No	**277**	**73.28**		**304**	**80.64**
p value			**0.0164**		
At age 30, live in rented house/ apartment					
Yes	156	41.27		137	36.34
No	222	58.73		240	63.66
p value			0.1645		

At age 30, type of area lived in					
Farm	4	1.06		5	1.32
Rural	34	8.99		30	7.94
Town	91	24.07		80	21.16
Suburb	137	36.24		138	36.51
Urban	112	29.63		125	33.07
p value			0.775		
At age 30, housing had indoor plumbing					
Yes	373	98.68		373	98.68
No	5	1.32		5	1.32
p value			1.0		
Age 30, household financial status (1-5: highest)					
1	7	1.85		4	1.06
2	47	12.43		43	11.41
3	210	55.56		213	56.50
4	90	23.81		93	24.67
5	24	6.35		24	6.37
p value			0.8998		
At age 30, how many others shared bedroom					
0	82	21.69		72	19.05
1	283	74.87		292	77.25
2+	13	3.44		14	3.70
p value			0.6612		
At age 30, how many cats/dogs/furry pets lived in the home					
Yes	34	8.99		31	8.20
No	344	91.01		347	91.80
p value			0.6971		
At age 30, house was within 1/2 mile of barns and stables where livestock were kept					
Yes	**79**	**20.95**		**51**	**13.53**
No	**298**	**79.05**		**326**	**86.47**
p value			**0.0069**		
In reference year, lived in rented house/ apartment					
Yes	54	14.29		61	16.18
No	324	85.71		316	83.82
p value			0.4688		
In reference year, type of area lived in					
Farm	2	0.53		2	0.53
Rural	45	11.97		45	11.94
Town	104	27.66		98	25.99
Suburb	160	42.55		157	41.64
Urban	65	17.29		75	19.89
p value			0.9217		
In reference year, housing had indoor plumbing					
Yes	378	100.00		375	99.73
No	0	0.00		1	0.27
p value			0.3157		

In reference year, self-ranked household financial status					
1	2	0.53		2	0.53
2	25	6.61		28	7.43
3	118	31.22		128	33.95
4	175	46.30		151	40.05
5 (highest)	58	15.34		68	18.04
p value			0.5354		
In reference year, how many others shared bedroom					
0	132	34.92		121	32.10
1	245	64.81		256	67.90
2+	1	0.26		0	0.00
p value			0.4235		
In reference year, how many cats/dogs/furry pets lived in the home					
Yes	13	3.44		13	3.44
No	365	96.56		365	96.56
p value			1.0		
At reference year, house was within 1/2 mile of barns and stables where livestock were kept					
Yes	**94**	**24.87**		**68**	**17.99**
No	**284**	**75.13**		**310**	**82.01**
p value			**0.0212**		
Relative weight at age 12, compared to others the same age					
Below average	92	23.34		109	28.76
About average	227	60.05		223	58.84
Above average	59	15.61		47	12.40
p value			0.2429		
Relative height at age 12, compared to others same age					
Taller	137	36.24		138	36.41
The same	152	40.21		155	40.90
Shorter	89	23.54		85	22.43
Don't know	0	0.00		1	0.26
p value			0.7714		
Body mass index at reference year					
Less than 25	187	49.73		183	48.93
25-29	106	28.19		114	30.48
30 or more	83	22.07		77	20.59
p value			0.7581		
Ever knowingly had infection with parasites					
Yes	37	9.84		41	11.02
No	339	90.16		331	88.98
p value			0.5972		
Ever had chicken pox infection (not vaccine)					
Yes	329	88.20		334	90.76
No	44	11.80		34	9.24
p value			0.2568		

Ever had measles infection (not vaccine)					
Yes	304	83.06		305	84.49
No	62	16.94		56	15.51
p value			0.6018		
Ever had mononucleosis (or kissing disease)					
Yes	**61**	**16.22**		**43**	**11.35**
No	**315**	**83.78**		**336**	**88.65**
p value			**0.0518**		
Ever had shingles or varicella zoster					
Yes	43	11.38		51	13.49
No	335	88.62		327	86.51
p value			0.3779		
Ever had mastitis (particularly while nursing)					
Yes	39	10.34		40	10.67
No	338	89.66		335	89.33
p value			0.8856		
Ever had chlamydia					
Yes	20	5.31		18	4.79
No	357	94.69		358	95.21
p value			0.7455		
Ever had urinary tract infection > 5 times					
Yes	86	22.81		87	23.14
No	291	77.19		289	76.86
p value			0.9152		
Ever had more than 5 vaginal yeast infections					
Yes	87	23.02		91	24.14
No	291	76.98		286	75.86
p value			0.7165		
Ever had genital herpes					
Yes	35	9.31		31	8.18
No	341	90.69		348	91.82
p value			0.5829		
Ever been told had stomach ulcer caused by Helicobacter pylori					
Yes	12	3.19		12	3.19
No	364	96.81		364	96.81
p value			1.0		
Ever had lyme disease					
Yes	4	1.06		3	0.79
No	374	98.94		376	99.21
p value			0.7015		
Ever had diabetes requiring insulin					
Yes	11	2.91		10	2.64
No	367	97.09		369	97.36

p value			0.8201		
Ever had severe gum disease requiring gum surgery					
Yes	46	12.17		46	12.14
No	332	87.83		333	87.86
p value			0.9892		
Age first diagnosed with mononucleosis					
Less than 16 yrs	24	39.34		15	34.88
17-19 yrs	19	31.15		14	32.56
20-24 yrs	9	14.75		8	18.60
25+ yrs	9	14.75		6	13.95
p value			0.9424		
Age at chicken pox					
Less than 5 yrs	47	12.43		46	12.14
5 yrs	66	17.46		52	13.72
6 yrs	57	15.08		65	17.15
7 yrs	56	14.81		64	16.89
8+ yrs	101	26.72		99	26.12
Unknown	51	13.49		53	13.98
p value			0.7328		
Age at measles					
No	62	16.40		56	14.78
LT 5 yrs	31	8.20		27	7.12
5 yrs	28	7.41		36	9.50
6 yrs	63	16.67		44	11.61
7 yrs	56	14.81		63	16.62
8+ yrs	124	32.80		126	33.25
Unknown	14	3.70		27	7.12
p value			0.1472		
Had a job requiring teaching or taking care of kids under age 5					
No teaching	216	57.14		240	63.32
Teach older kids	75	19.84		74	19.53
Teach kids < age 5	87	23.02		65	17.15
p value			0.1079		
Total number of C-sections in life					
0	334	88.36		327	86.28
1+	44	11.64		52	13.72
p value			0.3898		
Total kids raised (biological, adopted, grandkids, other)					
0	59	15.61		50	13.19
1	62	16.40		56	14.78
2	125	33.07		127	33.51
3	69	18.25		83	21.90
4+	63	16.67		63	16.62
p value			0.6713		
Ever adopt any children					
Yes	16	4.23		21	5.57
No	362	95.77		356	94.43
p value			0.3947		

Have any grandchildren					
Yes	192	50.79		203	53.70
No	186	49.21		175	46.30
p value			0.4232		
Ever raise other children not biologically yours					
Yes	41	10.85		42	11.11
No	337	89.15		336	88.89
p value			0.9074		
Ever work as teacher or professional daycare provider for 1 year					
Yes	162	42.86		139	36.97
No	216	57.14		237	63.03
p value			0.0987		
Ever work with children under age 5					
Yes	87	53.05		65	46.10
No	77	46.95		76	53.90
p value			0.2262		
Age at first menstrual period					
<12 yrs	83	22.13		86	22.81
12 yrs	128	34.13		125	33.16
13 yrs	94	25.07		104	27.59
14+ yrs	70	18.67		62	16.45
p value			0.7835		
Number of pregnancies					
None	48	12.70		51	13.46
1 pregnancy	46	12.17		39	10.29
2 pregnancies	91	24.07		91	24.01
3 pregnancies	71	18.78		84	22.16
4+ pregnancies	122	32.28		114	30.08
p value			0.7307		
Ever took birth control pills					
Yes	275	72.75		265	69.92
No	103	27.25		114	30.08
p value			0.3892		
Ever used prescription hormone therapy for menopause symptoms					
Yes	**193**	**51.06**		**226**	**59.63**
No	**185**	**48.94**		**153**	**40.37**
p value			**0.0177**		
Still having regular periods					
Yes	**23**	**8.65**		**4**	**1.19**
No	**243**	**91.35**		**332**	**96.81**
p value			**<0.0001**		
Had hysterectomy					
Yes	122	32.71		114	30.32
No	251	67.29		262	69.68
p value			0.4817		
Had ovaries removed					
Yes	81	21.43		85	22.67

No	297	78.57		290	77.33
p value			0.682		
Ever had tonsils removed					
Yes	200	52.91		204	53.83
No	178	47.09		175	46.17
p value			0.8006		
Ever had appendix removed					
Yes	80	21.28		91	24.01
No	296	78.72		288	75.99
p value			0.3695		
Ever had spleen removed					
Yes	1	0.26		0	0.00
No	377	99.74		379	100.00
p value					
In the last five years, ever took antibiotics for infection					
Yes	279	73.81		275	73.92
No	99	26.19		97	26.08
p value			0.9714		
Mother had breast cancer					
No	333	88.10		331	87.34
Yes	45	11.90		48	12.66
p value			0.7501		
Currently drinks 10 or more alcoholic beverages per week					
0	339	89.68		332	87.60
1	39	10.32		47	12.40
p value			0.3664		
Menopausal hormone replacement therapy use in last five years					
None/ Not last 5 yrs	**266**	**71.89**		**238**	**63.64**
Combined Estrogen/ progestin	**48**	**12.97**		**76**	**20.32**
Estrogen only	**56**	**15.14**		**60**	**16.04**
p value			**0.0184**		

Table 2: Odds ratios (OR) for breast cancer for suggestively associated proxy measures of microbial exposures as ascertained from cases (n=379) and controls (n=378), Northern California, 2006-08.

	Unadjusted*		Adjusted†	
	OR	95% CI	OR	95% CI
Attended preschool				
Yes	1		1	
No	1.47	0.93-2.31	1.47	0.93-2.34
Attended kindergarten				
Yes	**1**		**1**	
No	**1.55**	**1.09-2.22**	**1.52**	**1.06-2.18**
Before 18, attended boarding school where lived in dorm				
Yes	**1**		**1**	
No	**1.87**	**1.10-3.19**	**1.91**	**1.11-3.29**
At 6 months, lived within 1/2 mile of barns and stables where livestock were kept				
Yes	1		1	
No	1.01	0.71-1.45	1.05	0.73-1.51
At age 5, lived within 1/2 mile of barns and stables where livestock were kept				
Yes	1		1	
No	0.93	0.66-1.30	0.96	0.68-1.37
At age 12, lived within 1/2 mile of barns and stables where livestock were kept				
Yes	**1**		**1**	
No	**1.47**	**1.04-2.08**	**1.54**	**1.08-2.19**
At age 30, lived within 1/2 mile of barns and stables where livestock were kept				
Yes	**1**		**1**	
No	**1.64**	**1.11-2.41**	**1.57**	**1.05-2.33**
At reference year, lived within 1/2 mile of barns and stables where livestock were kept				
Yes	**1**		1	
No	**1.43**	**1.00-2.04**	1.42	0.98-2.04

Average number of days/month in last year reported taking lactobacillus or probiotic supplements				
None	1		1	
1+ days/mo	0.58	0.31-1.10	0.66	0.34-1.27

* Adjusted for age and race/ethnicity
† Adjusted for age, race, and breast cancer risk factors (first degree family history of breast cancer, current consumption of 10 or more alcoholic beverages per week, and current hormone replacement therapy use (none, current estrogen only, current combined EP)

Table 3: Multivariately adjusted odds ratios (OR) for breast cancer and 95% confidence intervals (CI) for suggestively associated proxy measures of microbial exposures considered separately by race/ethnicity of participants, Northern California, 2006-08.

	White, non-Hispanic*		Non-White*	
	OR	95% CI	OR	95% CI
Attended preschool				
Yes	1		1	
No	1.55	0.93-2.58	0.96	0.31-3.00
Attended kindergarten				
Yes	1		1	
No	1.43	0.95-2.16	1.69	0.82-3.49
Before 18, attended boarding school where lived in dorm				
Yes	**1**		1	
No	**2.13**	**1.16-3.90**	1.05	0.30-3.70
At age 6 months, lived within 1/2 mile of barns and stables where livestock were kept				
Yes	1		1	
No	0.96	0.63-1.48	1.36	0.66-2.80
At age 5, lived within 1/2 mile of barns and stables where livestock were kept				
Yes	1		1	
No	0.97	0.64-1.45	1.04	0.48-2.22
At age 12, lived within 1/2 mile of barns and stables where livestock were kept				
Yes	**1**		1	
No	**1.56**	**1.05-2.32**	1.61	0.72-3.59
At age 30, lived within 1/2 mile of barns and stables where livestock were kept				
Yes	**1**		1	
No	**1.52**	**0.99-2.34**	1.82	0.61-5.44
At reference year, lived within 1/2 mile of barns and stables where livestock were kept				
Yes	1		1	
No	1.37	0.93-2.01	2.05	0.62-6.79
Average number of days/month in last year reported taking lactobacillus or probiotic supplements				
None	1		1	
1+ days/mo	0.62	0.31-1.24	1.47	0.18-12.08

* Adjusted for age, specific ethnicity, and breast cancer risk factors (first degree family history of breast cancer, current consumption of 10 or more alcoholic beverages per week, and current hormone replacement therapy use (none, current estrogen only, current combined EP)

Table 4: Multivariately adjusted odds ratios (OR) for breast cancer and 95% confidence intervals (CI) for suggestively associated proxy measures of microbial exposures considered only for estrogen-receptor positive breast cancer (n=364), Northern California, 2006-08.

	Unadjusted*		Adjusted†	
	OR	95% CI	OR	95% CI
Attended preschool				
Yes	1		1	
No	1.50	0.93-2.43	1.51	0.92-2.47
Attended kindergarten				
Yes	**1**		**1**	
No	**1.56**	**1.08-2.26**	**1.55**	**1.07-2.26**
Before 18, attended boarding school where lived in dorm				
Yes	1		1	
No	1.67	0.97-2.90	1.69	0.97-2.95
At age 6 months, lived within 1/2 mile of barns and stables where livestock were kept				
Yes	1		1	
No	1.04	0.72-1.52	1.08	0.73-1.58
At age 5, lived within 1/2 mile of barns and stables where livestock were kept				
Yes	1		1	
No	0.94	0.65-1.34	0.98	0.67-1.42
At age 12, lived within 1/2 mile of barns and stables where livestock were kept				
Yes	**1**		**1**	
No	**1.51**	**1.05-2.18**	**1.61**	**1.10-2.35**
At age 30, lived within 1/2 mile of barns and stables where livestock were kept				
Yes	**1**		**1**	
No	**1.72**	**1.14-2.61**	**1.64**	**1.07-2.51**
At reference year, lived within 1/2 mile of barns and stables where livestock were kept				
Yes	**1**		**1**	
No	**1.31**	**0.90-1.91**	**1.29**	**0.88-1.89**
Average number of days/month in last year reported taking lactobacillus or probiotic supplements				
None	1		1	
Lacto 1+ days/mos	0.66	0.34-1.26	0.75	0.39-1.47

* Adjusted for age and race/ethnicity

† Adjusted for age, race, and breast cancer risk factors (first degree family history of breast cancer, current consumption of 10 or more alcoholic beverages per week, and current hormone replacement therapy use (none, current estrogen only, current combined EP).

Table 5: Odds ratios for neighborhood characteristics and breast cancer among respondents in the study (379 cases, 378 controls) and the larger pool of presumably eligible subjects (780 cases, 3550 control households), including calculation of % difference between groups.

	Respondents only			**All Subjects**			
Neighborhood characteristic	OR	95% CI		OR	95% CI		Bias (%)†
SES (quintiles							
1- Low	1.00			1.00			
2	3.38	0.95	12.01	1.99	1.05	3.74	41
3	1.97	0.62	6.24	2.24	1.24	4.07	-14
4	1.99	0.67	5.96	3.16	1.78	5.58	-59
5 - High	2.22	0.76	6.51	3.83	2.18	6.71	-72
Percent below 200% poverty							
1	1.24	0.79	1.96	2.11	1.67	2.65	-70
2	0.93	0.57	1.51	1.63	1.28	2.07	-75
3	0.94	0.58	1.54	1.31	1.02	1.68	-39
4	1.00			1.00			
Proportion 16+ yr olds with blue collar job							
1	0.82	0.54	1.27	1.58	1.27	1.97	-93
2	1.07	0.68	1.68	1.28	1.02	1.61	-20
3	0.76	0.47	1.23	0.99	0.78	1.26	-30
4	1.00			1.00			
Proportion of 25+ yr olds with college degrees							
1	1.00			1.00			
2	1.09	0.67	1.77	1.34	1.06	1.71	-23
3	0.99	0.62	1.58	1.50	1.18	1.89	-52
4	0.93	0.59	1.46	1.81	1.44	2.28	-95
Proportion of 25+ yr olds with out high school diploma							
1	0.88	0.56	1.39	1.88	1.48	2.39	-114
2	1.02	0.64	1.62	1.72	1.35	2.19	-69
3	1.68	1.01	2.78	1.73	1.36	2.21	-3
4	1.00			1.00			
Median household income							
1	1.00			1.00			
2	1.37	0.83	2.24	1.46	1.14	1.88	-7
3	1.39	0.86	2.26	1.90	1.49	2.41	-37
4	1.28	0.80	2.03	2.02	1.59	2.56	-58
Median gross rent							
1	1.00			1.00			
2	1.42	0.89	2.26	1.26	0.99	1.59	11
3	1.18	0.76	1.84	1.33	1.05	1.69	-13
4	1.12	0.74	1.69	1.85	1.47	2.31	-65
Median value of owner							

occupied houses							
1	1.00			1.00			
2	0.92	0.56	1.52	1.27	0.99	1.62	-38
3	1.03	0.66	1.60	1.61	1.27	2.03	-56
4	0.98	0.64	1.52	1.78	1.42	2.24	-82
Proportion 16+ yr olds unemployed							
1	1.21	0.79	1.85	1.75	1.40	2.20	-45
2	1.45	0.92	2.29	1.46	1.16	1.85	-1
3	1.00	0.63	1.58	1.28	1.01	1.62	-28
4	1.00			1.00			

† calculated as [OR(respondents)-OR(all subjects)]/OR(respondents)} x 100